NAVAL SUBMARINE MEDICAL RESEARCH LABORATORY

SUBMARINE BASE, GROTON, CONN.

REPORT NUMBER 1093

PERFORMANCE AND PREFERENCE WITH VARIOUS VDT PHOSPHORS

by

S. M. LURIA, David F. NERI and Christine SCHLICHTING

Naval Medical Research and Development Comm.
Research Work Unit M0100.001 1002

Released by:

C. A. Harvey, CAPT, MC, USN
Commanding Officer
Naval Submarine Medical Research Laboratory

24 April 1987

PERFORMANCE AND PREFERENCE WITH VARIOUS VDT PHOSPHORS

by

S. M. LURIA, David F. NERI, and Christine SCHLICHTING

NAVAL SUBMARINE MEDICAL RESEARCH LABORATORY

REPORT NUMBER 1093

Naval Medical Research and Development Command
Research Work Unit M0100.001-1002

Approved and Released by:

C. A. Harvey

C. A. HARVEY, CAPT, MC, USN
Commanding Officer
NAVSUBMEDRSCHLAB

SUMMARY PAGE

Problem

To determine the correlation between the preferences of VDT operators for variously colored phosphors and their performance on a visual search task with these phosphors.

Findings

Although many VDT operators have marked preferences and dislikes for various VDT phosphors, there is no significant relationship between these preferences and their performance in a difficult scanning task using the phosphors.

Application

Operator satisfaction and morale may be affected by the imposition of a given chromatic VDT phosphor, but performance is unlikely to be unaffected, at least on short-term tasks where scanning is involved.

Administrative Information

This investigation was conducted as part of Naval Medical Research and Development Command Work Unit M0100.001-1002. was submitted for review on 13 March 1987, approved for publication on 24 April 1987, and has been designated as Naval Submarine Medical Research Laboratory Report No. 1093.

Abstract

Subjects searched for target letters and symbols in VDT displays produced with phosphors of different colors: green, amber, yellow, red, blue, and white. There were no significant differences in search time, but there were significant differences in the number of errors made with different phosphors. Yellow produced the fewest errors and white the most. Speed and accuracy did not correlate with the preference ratings for the various phosphors.

PERFORMANCE AND PREFERENCE WITH VARIOUS VDT PHOSPHORS

S.M. Luria, David F. Neri, and Christine Schlichting

Monochromatic video display terminals (VDTs) are manufactured with different colored phosphors. For VDTs used primarily to display text, the two most widely found colors are green and amber. Many individuals have marked preferences for one of these colors, but it is not completely clear if performance is better with a preferred phosphor. Most of the studies in this area have compared the effects of different colors on performance (c.f., Carter and Carter, 1982), or have compared monochromatic with chromatic VDTs (c.f., McLean, 1965); some studies have compared older VDTs with those having such new features as the "shadow-mask" color displays (Zlotnick, 1986).

There have been many studies of the perceptibility of different colors and color-combinations. Typically it has been found that for a given background color, the opponent-color target was most visible (Ohlsson et al., 1981; Neri et al., 1986). Such differences in target visibility were more pronounced for some background colors than others. Both Ohlsson et al (1981) and Neri et al. (1986) found the smallest differences with a blue background and the largest with a reddish background. Neri et al. (1985) also compared target detection on a blue and a black background and found that, when the color contrast of the various targets was equated, there were no differences in target detection.

It is easy to imagine that preference has a significant effect on performance. Kroemer (1987, p. 6) has commented that "Personal likes and dislikes are very important, perhaps more so than what the "experts" say is appropriate." Yet the operator's preference for a given phosphor color may not produce the best performance. Smith (1986) has remarked that "colours that may appeal only to an individual's aesthetic judgment, but that ignore psycnophysics and task requirements, can degrade visual performance and comfort." The quality of a retinal image depends, of course, on the accuracy of focus (Ogle, 1961). Different colors have different focal lengths; blue has the shortest and red the longest. More accommodation is required to focus red light, and this becomes more difficult with age. Red light thus tends to produce more fatigue (Kinney et al., 1983), particularly with older people. On the other hand, it is widely conceded that acuity is poor with blue (Brindley, 1953, 1954), and extended attempts to read with that color might then also be fatiguing.

In this study, subjects carried out a rather difficult scanning task. Their speed and accuracy in searching for target letters in a text displayed by various phosphors were measured. These results were correlated with their preference ratings for the phosphors.

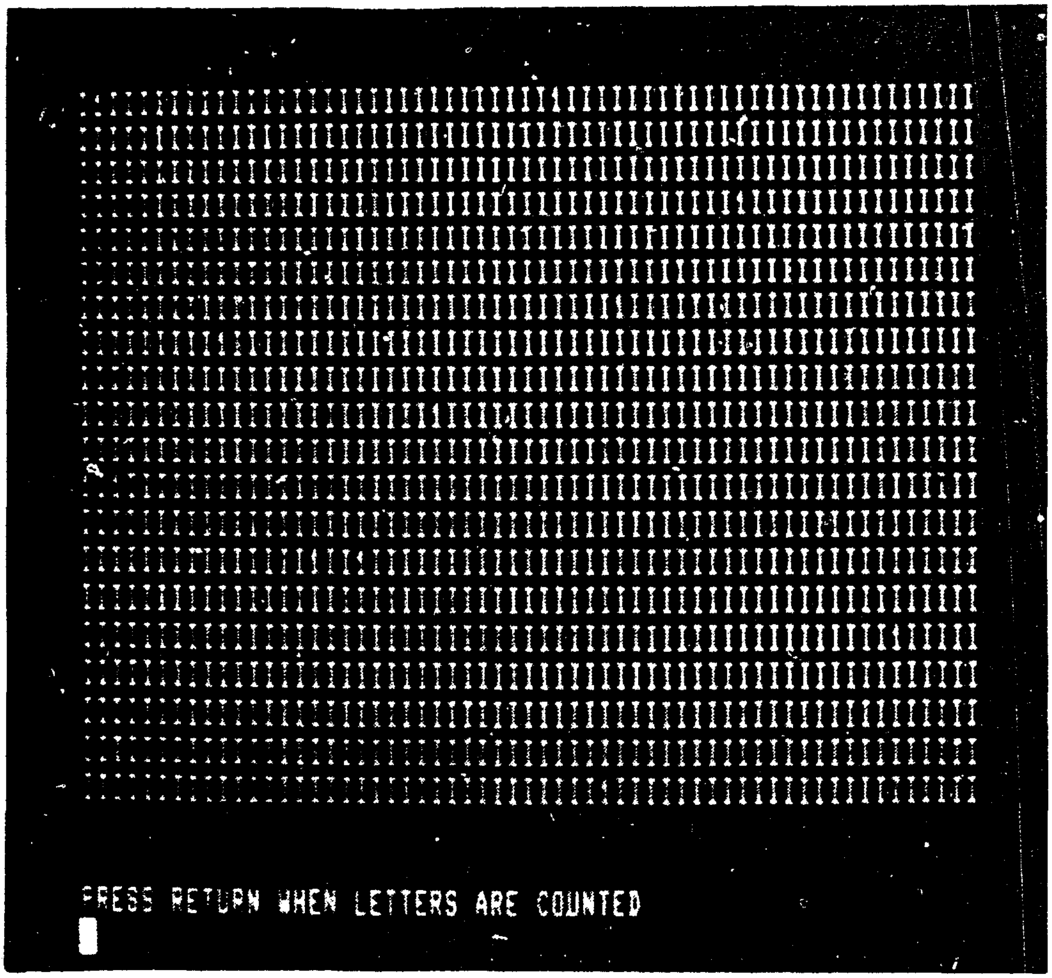

Figure 1. Photograph of the display with a background composed of the letter I and the number 1 as the target. There are 12 targets in this array.

EXPERIMENT I

METHOD

Subjects

Twenty staff members of the laboratory were subjects. They ranged in age from 20 to 57. They were not screened for refractive error or color vision. Those with eyeglasses wore them during the experiment. Several were known to be color defectives, which did not, hov ver, preclude their having firm preferences for one ph or the other.

Ta

subjects' task was to count as quickly and accurately as possible the number of times a target letter appeared in a background array consisting of another letter (Figure 1), and then to rate the desirability of the phosphor with which the display was presented. The array consisted of 20 lines, each line 60 letters in length. The computer inserted from 4 to 33 target letters randomly throughout the display, except that there were no more than three target letters in one line. The mean number of targets presented in a display was 16.5 $\pm$ 2.5. Each subject repeated this search 20 times, each time for a different target letter in a different array. The combinations were chosen to be difficult. One target was the number "0" with a background of the letter "O". The distinction in the computer font was quite small. Another target was the number "1" with a background of the letter "I", again a very slight difference (shown in Fig. 1). Table 1 gives the entire series of target-background combinations in the order in which they were presented to every subject. The number of target letters varied considerably and randomly for each presentation.

TABLE I

The series of target and background letters

Target/Bkgd	Target/Bkgd	Target/Bkgd	Target/Bkgd
1. c/e	6. U/O	11. I/J	16. e/o
2. 0/O	7. I/T	12. u/n	17. E/F
3. R/P	8. n/m	13. O/G	18. F/P
4. c/o	9. Q/O	14. g/q	19. G/C
5. 1/I	10. y/g	15. $/&	20. h/b

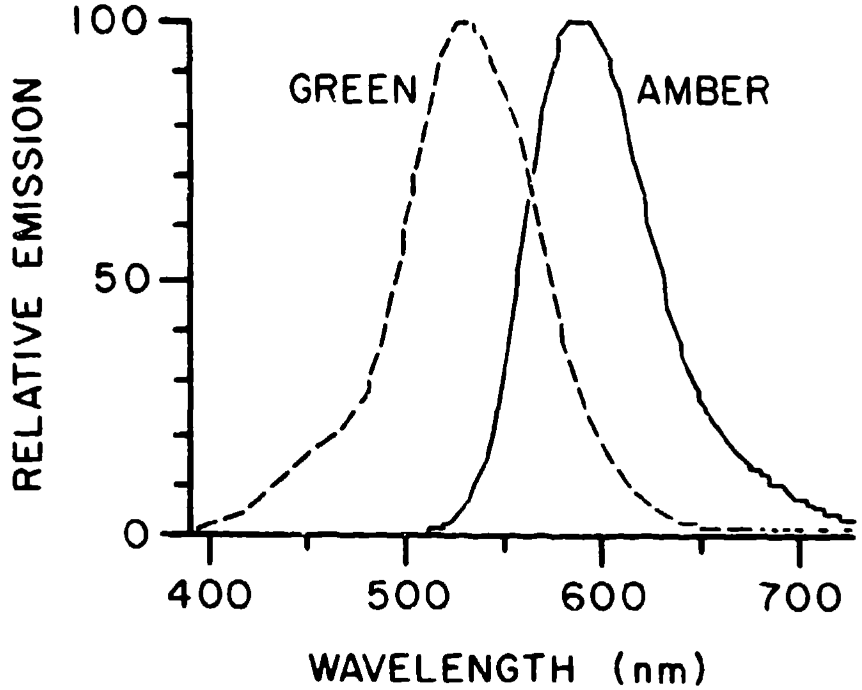

Figure 2. Spectral distributions of the phosphors in Experiment I.

Apparatus

The displays were presented on Micro-Term Model 301 terminals, one with a green and one with an amber phosphor. Their spectral distributions are shown in Figure 2. The screens measured 20 x 25 cm; the display itself was 12 x 15 cm. The subject positioned himself at a comfortable distance from the monitor; the viewing distance was therefore somewhat variable. The luminance of the two phosphors was equated to 16.5 c/m2. The tests were run in an office with all but two of the most distant overhead fluorescent fixtures off thus reducing the ambient illumination on the screen to 1.6 c/m2 and eliminating glare. This light level is only slightly higher than that falling on displays in the control rooms of submarines when they are rigged for red (Benson et al, 1987; Luria and Kobus, 1986).

Procedure

Half the subjects were tested first with the green phosphor, half with the amber phosphor. After explaining the purpose of the experiment to the subject, he was given practice runs searching for a target letter not included in Table 1 until he felt comfortable with the following procedure.

The subject first pressed a key to signal the computer that he was ready. The display appeared, and a computer controlled timer started. When the subject had counted the target letters, he pressed a key to stop the timer and erase the display. He then entered the number of targets. He then went through the series of target letters. The subjects were permitted to pause between the target-trials (when, of course, the clock was not running) if they wished. After completing the series, the subject was asked to rate the phosphor on a scale of 1 (poor) to 10 (good) for general desirability. He was then given a 10 minute break, and the monitor was exchanged. The subject then went through the same series of target letters again in the same order using the other phosphor, after which he rated that one. A test session took between 30 and 60 minutes.

The length of time that the subjects worked with a given phosphor averaged no more than half an hour. The criticism might be raised that this is too short a time. However, it is quite comparable with exposure times in other studies. For example, even in an experiment designed to investigate reading fatigue, Nordqvist et al. (1986) had their subjects read texts for 15 minutes, followed by 5 minutes of performance tests. This is more typical than having subjects observe for as much as a day, as Tullis (1981) did.

RESULTS

Table II shows the mean total number of targets missed by each subject during the 20 scans with each phosphor, the mean time taken by each subject to complete the search for the each target, and the mean rating given to each phosphor. None of these differences was significant according to the Wilcoxon matched-pairs signed-ranks test.

TABLE II

Mean Ratings, Errors, and Search Times in Experiment I

	Green	Amber
Rating	6.80	5.95
Errors	58.25	55.45
Time(sec)	50.05	50.50

The correlations between these performance measures and the subjects' ratings were also not significant (Table III). The correlation between errors and the ratings were $r = -.22$ with green and $r = .07$ with amber; those between search time and the ratings were $r = .38$ for green and $r = -.24$ for amber.

TABLE III

Performance-Preference Correlations in Experiment I

	Green	Amber
Errors	-.22	.07
Time	.38	-.24

The error rate negatively correlated with search time both for green ($r = -.18$) and for yellow ($r = -.24$). There was thus a tendency for more errors to be made when the search time was shorter, but neither correlation was significant.

There was, however, a significant difference in the number of errors made with the first phosphor compared to the second phosphor, showing the effect of practice (Table IV). For the first phosphor, an average of 61.6 targets were missed, whereas with the second phosphor only 52.1 targets were missed. This difference was significant ($p < .05$) according to the Wilcoxon test. Interestingly, there was no significant improvement in the average time taken to complete the scans (51.65 vs 48.90 sec). The mean rating given to the first phosphor was 5.85 and the mean rating given to the second was 6.85; the difference was not significant.

TABLE IV

Mean Scores for Order of Presentation in Experiment I

	First	Second
Rating	5.85	6.85
Errors*	61.6	52.1
Time(sec)	51.65	48.90

* $p < .05$

It is clear that there were reliable differences in the difficulty of finding the different targets, for the correlation between the error scores for each of the 20 targets with the two phosphors was $r = .94$ ($p < .01$), and the correlation between the search times for each target with the two phosphors was $r = .95$ ($p < .01$).

Not surprisingly, therefore, when the scores for only the 5 or 10 most difficult targets were considered, there were still no significant differences between the two phosphors for either number of errors or search times.

DISCUSSION

Whatever preferences the subjects may have had for one phosphor over the other, there were no significant differences between the mean preference ratings. That is, for every subject who preferred green, another preferred amber. There was also no difference between the performance measures for the two phosphors. Nor was there a significant correlation between the subjects' preferences and their performance.

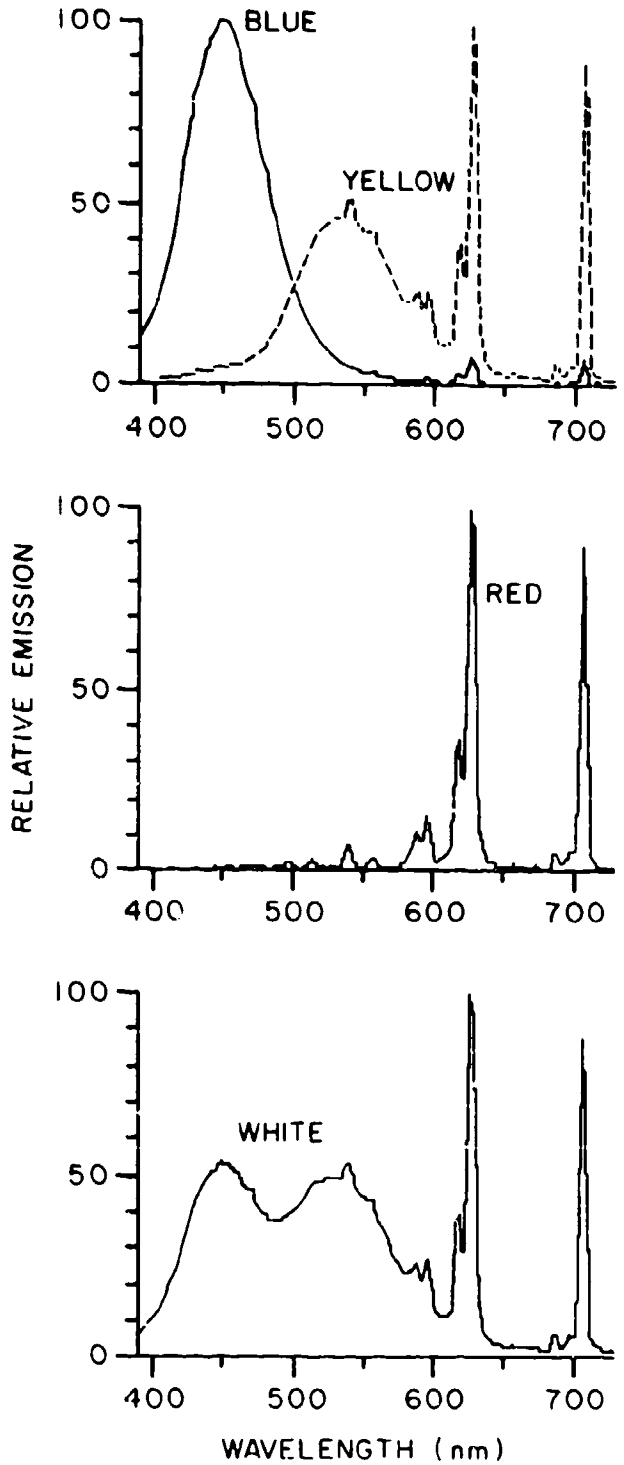

Figure 3. Spectral distributions of the phosphors in Experiment II.

EXPERIMENT II

Although it is unlikely that many other colors will be used for typical word processing, it is quite likely that a variety of colors will be used to present alphanumeric information in various types of displays which are found, for example, on submarines.

For that reason, the experiment was repeated with other phosphor colors, red, white, blue, and yellow.

METHOD

Subjects

Ten men, most of them enlisted men awaiting entrance to submarine school, volunteered to participate. Again no visual screening was carried out. Those who had refractive corrections wore their eyeglasses during the experiment.

Apparatus

The displays were presented on a Visual 241 Color Monitor, driven by the same computer as in Experiment I. Their spectral distributions are shown in Figure 3. The luminances of different phosphors are typically quite different and not adjustable, as Van Nes (1986) has pointed out. The luminances of our colors were adjustable, however, and rather than equate them as was done in Expt. 1, each subject was permitted to adjust each phosphor to the luminance level he desired.

Procedure

Each subject completed the same search task as in Experiment I for each of the four phosphors in one session. The subject was allowed to pause briefly between trials when he wished, and there were five minute pauses between phosphors. A session lasted 1 to 2 hours.

The phosphors were presented in a counterbalanced order to the first eight subjects; the presentations would have been completely counterbalanced had it not been necessary to discontinue the study after 10 subjects had been tested.

RESULTS

Table V gives the mean scores for the four phosphors. Yellow was the most preferred color with a mean rating of 6.7, followed closely by white (6.5). Blue was, not surprisingly, the least preferred color (4.85). However, one of the subjects rated blue most highly, and two others rated it second best, and thus these differences were not significant ($\underline{p} < .25$) according to the Friedman analysis of variance by ranks.

TABLE V

Mean Ratings, Errors, and Search Times
in Experiment II

	Red	White	Blue	Yellow
Rating	5.2	6.5	4.8	6.7
Errors*	35.9	38.9	32.4	30.8
Time(sec)	47.2	51.2	49.2	49.5

----
* $p < .001$

The difference in the mean number of errors made with the different phosphors was highly significant, according to the Friedman analysis ($p < .001$). The fewest number of errors was made with the yellow phosphor, but the next fewest was made with the blue. The highest number of errors occurred with the white phosphor.

There were no significant differences in the search times with the different phosphors ($p < .60$).

Table VI shows the correlations between the ratings and the error rate or the search time. Most were negative, showing a tendency for higher ratings to be give to those phosphors which resulted in fewer errors or shorter search times.

TABLE VI

Performance-Preference Correlations
in Experiment II

	Red	White	Blue	Yellow
Errors	-.12	-.05	.48	-.21
Time	-.58	-.19	-.28	.18

Table VII gives the correlations between both errors and time and the preference for the phosphors in the order in which they were presented. Only one reached statistical significance-- that between error rate and ratings for the third phosphor to be presented-- but it is most likely that this was simply chance.

TABLE VII

Performance-Preference Correlations for Order of Presentation

	First	Second	Third	Fourth
Errors	-.19	-.26	.71*	-.37
Time	-.17	.36	-.48	-.06

* $p < .05$

The results were again analyzed according to order of presentation (Table VIII). There were no significant differences in the mean ratings, but both mean errors ($p < .05$) and mean search times ($p < .02$) declined significantly from the first to the second presentation.

TABLE VIII

Mean Scores with Order of Presentation in Experiment II

	First	Second	Third	Fourth
Rating	4.8	7.4	5.3	5.8
Error*	39.5	35.3	33.8	29.4
Time(sec)**	55.2	47.6	46.7	47.6

* $p < .05$

** $p < .02$

DISCUSSION

Although there is a tendency for phosphors which result in fewer errors and shorter search times to be rated better, the correlations obtained in this study were not significant. Moreover, for every phosphor, the ratings varied virtually throughout the possible range. Blue, for example, received the lowest mean rating, but although one subject gave it a rating of 1.5, another gave it a rating of 9.5, the highest of the four he worked with. Yellow received the highest mean rating, yet the ratings ranged from 1 to 9.

The performance data are somewhat surprising. Blue has generally been believed to be the worst color for textual material since the earliest studies (cf., Hartridge, 1947, p. 595; Borish, 1975, p. 361). Stiles (1949) and Brindley (1954) showed a lower acuity for the blue color mechanism. On the other hand, several studies have reported no differences in acuity for different wavelengths if the luminances are matched (Shlaer, et al, 1942; Baker, 1949) and no differences in contrast sensitivity for different wavelengths (Campbell and Green, 1965; Van Nes and Bouman, 1967; Verona, 1978; Nelson and Halberg, 1979; Neri and Kinney, 1982). The present results showed significant differences in the number of errors with the various phosphors (Table V). The smallest number of errors was made with yellow. Although blue is generally considered to be the worst color for textual material-- and indeed it received the lowest mean rating-- its error rate was relatively low. White produced the highest number of errors and the longest mean search time. The same results were reported by Santucci et al. (1982). And Ohlsson et al. (1981) reported that white letters on a blue background produced the best legibility.

The reason for this cannot be an unfavorable difference in brightness, since the subjects were allowed to set the brightness level of each phosphor until it was comfortable for them. One reason may be that studies relating visual acuity or contrast sensitivity to wavelength have generally used threshold measurements as their dependent variable; studies testing legibility of displays deal with stimuli far above threshold. Moreover, the color of the background has a marked effect; Ohlsson et al (1981) found great differences in the ranges of scanning times as a function of the background color. When a colored background is used, two things happen. First, there are greater differences in the range of contrast values for the variously colored letters than with a neutral background. Second, chromatic aberration may be induced which produces acuity differences.

In this study, the background was black, and the subjects adjusted the luminance to their desired level Under these conditions, blue text is apparently not the most objectionable.

In sum, observers have quite different preferences, and these preferences do not correlate highly with performance.

ACKNOWLEDGMENT

We thank Bob Lane for writing the program to present these experiments.

REFERENCES

Baker, K.E. (1949). Some variables influencing vernier acuity. J. Opt. Soc. Am. 39: 567-576.

Benson, W., Ghirardi, L.F., Kobus, D.A., Lambert, R.E., Luria, S.M., Massey, R., Oswald, L.J., and Plath, D.W. (1978). Report on the Trident SCS illumination committee: Analysis and solution of illumination-related problems on the ship control station of Ohio class submarines. San Diego: Naval Personnel Research and Development Center Rept. (In press).

Borish, I.M. (1975). Clinical Refraction (3rd Ed.) Chicago: The Professional Press.

Brindley, G.S. (1953). The effects on colour vision of adaptation to very bright lights. J. Physiol. (London) 122: 332-350.

Brindley, G.S. (1954). The summation areas of human colour-receptive mechanism at increment threshold. J. Physiol. (London) 124: 400-408.

Campbell, F.W. and Green, D.G. (1965). Optical and retinal factors affecting visual resolution. J. Physiol. (London) 176: 576-593.

Carter, R.C. and Carter, E.C. (1982). Visual search with color. J. Exp. Psychol.: Hum. Percep. & Perform. 8:127-136.

Hartridge, H. (1947). The visual perception of fine detail. Phil. Trans. Royl Soc. London (B): 232:519-671.

Kinney, J.A.S., Neri, D.F., Mercado, D.T., and Ryan, A.P. (1983). Visual fatigue in sonar control rooms lighted by red, white or blue illumination. NSMRL Rep. No. 1000, Groton, CT: Naval Submarine Medical Research Laboratory.

Kroemer, K.H.E. (1987). Letter to the editor. Human Fact. Bull. 30:6.

Luria, S.M. and Kobus, D.A. (1985). Acceptability of low level white lighting in the control room at sea. NSMRL Rep. 1050. Groton, CT: Naval Submarine Medical Research Laboratory.

McLean, M.V. (1965). Brightness contrast, color contrast, and legibility. Human Factors 7: 521-526.

Nelson, M.A. and Halberg, R.L. (1979). Visual contrast sensitivity functions obtained with colored and achromatic gratings. Hum. Fact. 21:225-228.

Neri, D.F. and Kinney, J.A.S. (1982) Contrast sensitivity measured in low levels of red, white, and blue ambient illumination. NSMRL Rep. No., 989. Groton, CT: Naval Submarine Medical Research Laboratory.

Neri, D.F., Kobus, D.A., Luria, S.M., Fan, S., and Zannelli, D. (1985). Effect of background/foreground color coding on detection in acoustic data displays. NSMRL Rep. No. 1040/NUSC Tech. Rep. 7325. Groton, CT: Naval Submarine Medical Research Laboratory.

Neri, D.F., S.M. Luria, and D.A. Kobus (1986). The detection of various color combinations under different chromatic ambient illuminations. Aviat. Space & Environ. Med. 57:555-560.

Nordqvist, T., Ohlsson, K., and Nilsson, L-G. (1986). Fatigue and reading of text on videotex. Human Factors 28: 353-363.

Ogle, K.N. (1961). Foveal contrast thresholds with blurring of the retinal image and increasing size of test stimulus. J. Opt. Soc. Am. 51:862-869.

Ohlsson, K., Nilsson, L-G., and Ronnberg, J. (1981). Speed and accuracy in scanning as a function of combinations of text and background colors. Man-Machine Studies 14:215-223.

Santucci, G., Menu, J.P., and Valot, C. (1982). Visual acuity in color contrast on cathode ray tubes: role of luminance, hue, and saturation contrasts. Aviat. Space & Environ. Med. 53: 478-484.

Shlaer, S., Smith, E.L., and Chase, A.M. (1942). Visual acuity and illumination in different spectral regions. J. Gen. Physiol. 25: 553-569.

Smith, W. (1986). Computer colour: psychophysics, task application, and aesthetics. International Scientific Conference: Work with Display Units, Stockholm, May 12-15.

Stiles, W.S. (1949). Increment thresholds and the mechanism of colour vision. Doc. Ophthalmol. 3: 138-163.

Tullis, T.S. (1981). An evaluation of alphanumeric, graphic, and color information displays. Human Factors 23: 541-550.

Van Nes, F.L. (1986). Space, colour and typography on visual display terminals. Behav. and Info. Tech. 5: 99-118.

Van Nes, F.L. and Bouman, M.A. (1967). Spatial modulation transfer in the human eye. J. Opt. Soc. Am. 57: 401-406.

Verona, R.W. (1978). Contrast sensitivity of the human eye to various display phosphor types. Paper presented at the Society for Information Display Meeting, San Francisco.

Zlotnick, M.A. (1986). Effects of shadow mask color vs. monochrome CRTs on display and interpretation of acoustic information. Burbank, CA: Lockheed-California Co., 1986,

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: NSMRL Rpt #1093 — 2. GOVT ACCESSION NO.: ADA184085	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): PERFORMANCE AND PREFERENCE WITH VARIOUS VDT PHOSPHORS	5. TYPE OF REPORT & PERIOD COVERED: Interim
	6. PERFORMING ORG. REPORT NUMBER: No. 1093
7. AUTHOR(s): S. M. LURIA, David F. NERI, and Christine SCHLICHTING	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Naval Submarine Medical Research Laboratory, Naval Submarine Base New London, Groton, CT 06349-5900	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: M0100.001-1002
11. CONTROLLING OFFICE NAME AND ADDRESS: Naval Medical Research & Development Command, Naval Medical Command, National Capital Region, Bethesda, MD 20814-5044	12. REPORT DATE: 24 Apr 1987
	13. NUMBER OF PAGES: 15
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

video display terminals; phosphors;

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Subjects searched for target letters and symbols in VDT displays produced with phosphors of different colors: green, amber, yellow, red, blue, and white. There were no significant differences in search time, but there were significant differences in the number of errors made with different phosphors. Yellow produced the fewest errors and white the most. Speed and accuracy did not correlate with the preference ratings for the various phosphors.

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)